L2 ANSWER 23 OF 23 REGISTRY COPYRIGHT 2002 ACS
RN 148514-28-7 REGISTRY
CN DNA (Mycobacterium tuberculosis mammalian cell invasion protein gene fragment) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Deoxyribonucleic acid (Mycobacterium tuberculosis mammalian cell invasion protein gene fragment)
SQL 1535

SEQ 1051 agcaaccagc aatacgacgg catgtcacgg ctaagtggct acctgacccc
==== ========== ====
HITS AT: 1057-1074

FILE 'CAPLUS' ENTERED AT 10:45:56 ON 03 APR 2002
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

FILE COVERS 1907 - 3 Apr 2002 VOL 136 ISS 14
FILE LAST UPDATED: 2 Apr 2002 (20020402/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

CAS roles have been modified effective December 16, 2001. Please check your SDI profiles to see if they need to be revised. For information on CAS roles, enter HELP ROLES at an arrow prompt or use the CAS Roles thesaurus (/RL field) in this file.

The P indicator for Preparations was not generated for all of the CAS Registry Numbers that were added to the CAS files between 12/27/01 and 1/23/02. As of 1/23/02, the situation has been resolved. Searches and/or SDIs in the H/Z/CA/CAplus files incorporating CAS Registry Numbers with the P indicator executed between 12/27/01 and 1/23/02 may be incomplete. See the NEWS message on this topic for more information.

L3 9 L2

=> d ibib abs hitrn l3 1-9; fil hom

L3 ANSWER 1 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2001:703734 CAPLUS
DOCUMENT NUMBER: 135:268261
TITLE: DNA sequences for strain analysis in Mycobacterium tuberculosis
INVENTOR(S): Fleischmann, Robert David; White, Owen Richardson; Fraser, Claire Marie; Venter, John Craig
PATENT ASSIGNEE(S): The Institute for Genomic Research, USA
SOURCE: U.S., 3 pp.

Searched by Barb O'Bryen, STIC 308-4291

CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6294328	B1	20010925	US 1998-103840	19980624

AB The present invention is directed to novel methodol. whereby different populations of the tuberculosis bacterial pathogen, Mycobacterium tuberculosis, or related Mycobacteria, can be genetically classified in relation to other isolates. DNA sequences for strains H37Rv (4,411,529 bp) and CDC 1551 (4,403,765 bp) are provided. Sites in the genome of Mycobacterium, which define previously unrecognized points of variability, are disclosed. The existence of this variability is of use to the clinician in order to consistently det. the identity of isolates of Mycobacterium responsible for individual cases of disease or disease outbreaks, thus suggesting appropriate choices for treatment protocols.

IT **362640-82-2 362724-95-6**
RL: BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses)
(nucleotide sequence; DNA sequences for strain anal. in Mycobacterium tuberculosis)

REFERENCE COUNT: 17 THERE ARE 17 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3 ANSWER 2 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2001:693805 CAPLUS
DOCUMENT NUMBER: 135:252751
TITLE: Detection of Mycobacterium tuberculosis by PCR amplification of REP13E12 repeated sequence
INVENTOR(S): Lee, Tae Yoon; Kim, Sung Kwang; Lee, Jong Seok; Lee, Jai Youl
PATENT ASSIGNEE(S): S. Korea
SOURCE: U.S. Pat. Appl. Publ., 12 pp.
CODEN: USXXCO
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2001023065	A1	20010920	US 2001-785904	20010216
CN 1310235	A	20010829	CN 2001-104597	20010219

PRIORITY APPLN. INFO.: KR 2000-7984 A 20000219

AB A method for detecting Mycobacterium tuberculosis by the polymerase chain reaction (PCR) amplification of the REP13E12 repeated sequence is provided. More particularly, a method for detecting Mycobacterium tuberculosis in clin. specimen by the PCR amplification of all or some of the REP13E12 repeated sequence is provided. REP13E12 is a new repeated sequence comprising 453 bp cloned from the division cell of Mycobacterium tuberculosis derived from Korea, and which only exists in the M. tuberculosis complex. The method shows excellent sensitivity and specificity.

IT **362070-37-9 362070-38-0**
RL: ARG (Analytical reagent use); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); USES (Uses)
(PCR primer; detection of Mycobacterium tuberculosis by PCR amplification of REP13E12 repeated sequence)

IT **362070-35-7 362070-36-8**
RL: ANT (Analyte); BOC (Biological occurrence); BSU (Biological study,

unclassified); PRP (Properties); ANST (Analytical study); BIOL (Biological study); OCCU (Occurrence)
(nucleotide sequence; detection of Mycobacterium tuberculosis by PCR amplification of REP13E12 repeated sequence)

L3 ANSWER 3 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 2001:31632 CAPLUS
DOCUMENT NUMBER: 134:111206
TITLE: Method of making and identifying attenuated microorganisms, compositions utilizing the sequences responsible for attenuation, and preparations containing attenuated microorganisms
INVENTOR(S): Gicquel, Brigitte; Guilhot, Christophe; Camacho, Luis
PATENT ASSIGNEE(S): Institut Pasteur, Fr.
SOURCE: PCT Int. Appl., 159 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001002555	A1	20010111	WO 2000-IB950	20000706

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.: US 1999-142982P P 19990706
US 1999-142833P P 19990708

AB A functional genomic approach for identification of mutants of microorganisms that are unable to grow under certain specific conditions is disclosed. In one aspect of the invention, a method is provided in which a library of signature tagged transposon mutants (STM) is constructed and screened for mutants attenuated in pathogenicity. The method is esp. useful for identifying loci involved in pathogenicity. The method is well suited to identification of mutant actinomycetales, such as mycobacteria. To perform an STM in M. tuberculosis, plasmid pCG113 was constructed, comprising a temp.-sensitive-sacB vector carrying an IS1096 deriv. with a unique restriction site permitting the insertion of DNA signature tags. This allows efficient counter-selection of the plasmid at 39.degree. on sucrose and isolation of large nos. of M. tuberculosis transposition mutants. The method is useful for, among other things, drug discovery and construction of vaccines.

IT **190550-78-8**
RL: PRP (Properties)
(unclaimed nucleotide sequence; method of making and identifying attenuated microorganisms, compns. utilizing the sequences responsible for attenuation, and prepns. contg. attenuated microorganisms)

REFERENCE COUNT: 14 THERE ARE 14 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3 ANSWER 4 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 1999:251841 CAPLUS
DOCUMENT NUMBER: 131:83939
TITLE: New insertion sequences and a novel repeated sequence in the genome of Mycobacterium tuberculosis H37Rv
AUTHOR(S): Gordon, Stephen V.; Heym, Beate; Parkhill, Julian;

Barrell, Bart; Cole, Stewart T.
CORPORATE SOURCE: Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, Paris, 75724, Fr.
SOURCE: Microbiology (Reading, U. K.) (1999), 145(4), 881-892 April
CODEN: MROBEO; ISSN: 1350-0872
PUBLISHER: Society for General Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The genome sequence of Mycobacterium tuberculosis H37Rv was found to contain 56 loci with homol. to insertion sequences (ISs). As well as the previously described IS6110, IS1081, IS1547 and IS-like elements, new ISs belonging to the IS3, IS5, IS21, IS30, IS110, IS256 and ISL3 families were identified. In addn., six ISs created a grouping of their own to form a new family (the IS1535 family). Elements with similarity to ISs in other actinomycetes were identified, suggesting the movement of ISs between related genera. The location of ISs on the chromosome revealed that an approx. 600 kb region close to the origin of replication lacks ISs, pointing to the possible detrimental effect of insertions in this area. Anal. of the distribution of ISs through the tubercle strains Mycobacterium africanum, M. microti, M. bovis, M. bovis BCG Pasteur, M. tuberculosis H37Ra, M. tuberculosis CSU#93 and 29 clin. isolates revealed that only IS1532, IS1533, IS1534, and IS1561' were absent from some of the strains tested. A novel repeated sequence, the REP13E12 family, is described that is present in seven copies on the M. tuberculosis H37Rv chromosome and which contains a probable phage attachment site. This study therefore offers an insight into the possible role of ISs and repetitive elements in the evolution of the M. tuberculosis genome, as well as identifying genetic markers that may be useful for phylogenetic and epidemiol. anal. of the tubercle complex.

IT **178353-85-0**, GenBank Z74410
RL: BOC (Biological occurrence); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)
(IS1605'-contg. fragment; new insertion sequences and novel repeated sequence in genome of Mycobacterium tuberculosis H37Rv)

IT **190741-74-3**, GenBank Z95586
RL: BOC (Biological occurrence); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)
(REP251-contg. fragment; new insertion sequences and novel repeated sequence in genome of Mycobacterium tuberculosis H37Rv)

IT **190550-78-8**, GenBank Z95390
RL: BOC (Biological occurrence); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)
(REP336-contg. fragment; new insertion sequences and novel repeated sequence in genome of Mycobacterium tuberculosis H37Rv)

REFERENCE COUNT: 36 THERE ARE 36 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3 ANSWER 5 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 1998:773005 CAPLUS
DOCUMENT NUMBER: 130:120325
TITLE: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. [Erratum to document cited in CA129:77224]
AUTHOR(S): Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. V.; Eiglmeier, K.; Gas, S.; Barry, C. E., III; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M. A.; Rajandream, M.-A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J. E.; Taylor, K.;

Whitehead, S.; Barrell, B. G.
CORPORATE SOURCE: Sanger Cent., Wellcome Trust Genome Campus, Hinxton, CB10 1SA, UK
SOURCE: Nature (London) (1998), 396(6707), 190-198
CODEN: NATUAS; ISSN: 0028-0836
PUBLISHER: Macmillan Magazines
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Table 1 was published with some symbols missing; the correct version can be found at http://www.sanger.ac.uk and is given here. In Fig. 2, Rv0649 was incorrectly labeled as fadD37 instead of fabD2. Two of the genes for mycolyl transferases were inverted: Rv0129c encodes antigen 85C and not 85c' as stated, whereas Rv3803c codes for the secreted protein MPT51 and not antigen 85C (Infect. Immun. 59, 372-382; 1991); Rv3803c is now designated fbpD. The sequence of Rv0746 from M. bovis BCG-Pasteur presented in Fig. 5 b was incorrect and should have shown a 16-codon deletion instead of 29.
IT **178353-85-0**, GenBank Z74410 **190550-78-8**, GenBank Z95391
190741-74-3, GenBank Z95586
RL: PRP (Properties)
(deciphering the biol. of Mycobacterium tuberculosis from the complete genome sequence (Erratum))

L3 ANSWER 6 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 1998:389708 CAPLUS
DOCUMENT NUMBER: 129:77224
TITLE: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence
AUTHOR(S): Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. V.; Eiglmeier, K.; Gas, S.; Barry, C. E., III.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M. A.; Rajandream, M.-A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G.
CORPORATE SOURCE: Sanger Cent., Wellcome Trust Genome Campus, Hinxton, CB10 1SA, UK
SOURCE: Nature (London) (1998), 393(6685), 537-544
CODEN: NATUAS; ISSN: 0028-0836
PUBLISHER: Macmillan Magazines
DOCUMENT TYPE: Journal
LANGUAGE: English
AB Countless millions of people have died from tuberculosis, a chronic infectious disease caused by the tubercle bacillus. The complete genome sequence of the best-characterized strain of Mycobacterium tuberculosis, H37Rv, was detd. and analyzed in order to improve our understanding of the biol. of this slow-growing pathogen and to help the conception of new prophylactic and therapeutic interventions. The genome comprises 4,411,529 base pairs, contains around 4000 genes, and has a very high G+C content that is reflected in the biased amino acid content of the proteins. M. tuberculosis differs radically from other bacteria in that a very large portion of its coding capacity is devoted to the prodn. of enzymes involved in lipogenesis and lipolysis, and to 2 new families of glycine-rich proteins with a repetitive structure that may represent a source of antigenic variation.
IT **178353-85-0**, GenBank Z74410 **190550-78-8**, GenBank Z95390
190741-74-3, GenBank Z95586
RL: PRP (Properties)
(nucleotide sequence; deciphering the biol. of Mycobacterium

tuberculosis from the complete genome sequence)

L3 ANSWER 7 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 1996:623197 CAPLUS
DOCUMENT NUMBER: 125:272760
TITLE: DNA molecule encoding for cellular uptake of Mycobacterium tuberculosis
INVENTOR(S): Riley, Lee W.
PATENT ASSIGNEE(S): Cornell Research Foundation, Inc., USA
SOURCE: PCT Int. Appl., 66 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 2
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9626275	A1	19960829	WO 1996-US2155	19960220
W: BR, CA, CN, JP, US				
RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
US 6008201	A	19991228	US 1995-464052	19950605
US 6214543	B1	20010410	US 1995-461002	19950605
CA 2212870	AA	19960829	CA 1996-2212870	19960220
EP 811066	A1	19971210	EP 1996-906506	19960220
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE				
CN 1182454	A	19980520	CN 1996-193436	19960220
BR 9607406	A	19980707	BR 1996-7406	19960220
JP 11500316	T2	19990112	JP 1996-525764	19960220

PRIORITY APPLN. INFO.: US 1995-392210 A 19950222
US 1993-118442 B2 19930902
WO 1996-US2155 W 19960220

AB A DNA mol. conferring on Mycobacterium tuberculosis an ability to enter mammalian cells and to survive within macrophages is isolated from M. tuberculosis. The DNA mol. contains 2 open reading frames; ORF-1 occurs at residues 181-807 of the 1535-nucleotide fragment and encodes a protein of 209 amino acids (22-28 kDa) which functions to mediate entry of M. tuberculosis into mammalian cells, whereas ORF-2 encodes a protein of 216 amino acids (>21 kDa). The protein(s) encoded by this gene fragment is useful in vaccines to prevent infection by Mycobacterium tuberculosis, while the antibodies raised against this protein can be employed n passively immunizing those already infected by the organism. Both these proteins and antibodies may be utilized in diagnostic assays to detect Mycobacterium tuberculosis in tissue or bodily fluids. The protein(s) can be assocd. with various other therapeutic materials, for administration to mammals, particularly humans, to achieve uptake of those materials by such cells.

IT **182239-95-8**
RL: ARG (Analytical reagent use); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); USES (Uses)
(amino acid sequence; DNA mol. encoding for cellular uptake of Mycobacterium tuberculosis)

IT **148514-28-7**
RL: ARG (Analytical reagent use); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); USES (Uses)
(nucleotide sequence; DNA mol. encoding for cellular uptake of Mycobacterium tuberculosis)

L3 ANSWER 8 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 1996:140980 CAPLUS
DOCUMENT NUMBER: 124:222277
TITLE: Molecular analysis of genetic differences between Mycobacterium bovis BCG and virulent M. bovis

AUTHOR(S): Mahairas, Gregory G.; Sabo, Peter J.; Hickey, Mark J.; Singh, Devinder C.; Stover, C. Kendall
CORPORATE SOURCE: Lab. Tuberculosis and Mol. Microbiol., PathoGenesis Corp., Seattle, WA, 98119, USA
SOURCE: J. Bacteriol. (1996), 178(5), 1274-82
CODEN: JOBAAY; ISSN: 0021-9193
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The live attenuated bacillus Calmette-Guerin (BCG) vaccine for the prevention of disease assocd. with Mycobacterium tuberculosis was derived from the closely related virulent tubercle bacillus, Mycobacterium bovis. Although the BCG vaccine has been one of the most widely used vaccines in the world for over 40 yr, the genetic basis of BCG's attenuation has never been elucidated. We employed subtractive genomic hybridization to identify genetic differences between virulent M. bovis and M. tuberculosis and avirulent BCG. Three distinct genomic regions of difference (designated RD1 to RD3) were found to be deleted from BCG, and the precise junctions and DNA sequence of each deletion were detd. RD3, a 9.3-kb genomic segment present in virulent lab. strains of M. bovis and M. tuberculosis, was absent from BCG and 84% of virulent clin. isolates. RD2, a 10.7-kb DNA segment contg. a novel repetitive element and the previously identified mpt-64 gene, was conserved in all virulent lab. and clin. tubercle bacilli tested and was deleted only from substrains derived from the original BCG Pasteur strain after 1925. Thus, the RD2 deletion occurred after the original derivation of BCG. RD1, a 9.5-kb DNA segment found to be deleted from all BCG substrains, was conserved in all virulent lab. and clin. isolates of M. bovis and M. tuberculosis tested. The reintroduction of RD1 into BCG repressed the expression of at least 10 proteins and resulted in a protein expression profile almost identical to that of virulent M. bovis and M. tuberculosis, as detd. by two-dimensional gel electrophoresis. These data indicate a role for RD1 in the regulation of multiple genetic loci, suggesting that the loss of virulence by BCG is due to a regulatory mutation. These findings may be applicable to the rational design of a new attenuated tuberculosis vaccine and the development of new diagnostic tests to distinguish BCG vaccination from tuberculosis infection.

IT **170316-51-5**, GenBank U35017 **170316-55-9**, GenBank U35021
RL: BOC (Biological occurrence); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)
(nucleotide sequence; mol. anal. of genetic differences between Mycobacterium bovis BCG and virulent Mycobacterium bovis)

L3 ANSWER 9 OF 9 CAPLUS COPYRIGHT 2002 ACS
ACCESSION NUMBER: 1995:682717 CAPLUS
DOCUMENT NUMBER: 123:80984
TITLE: Molecular cloning of gene for cellular uptake of Mycobacterium tuberculosis
INVENTOR(S): Riley, Lee W.
PATENT ASSIGNEE(S): Cornell Research Foundation, Inc., USA
SOURCE: PCT Int. Appl., 46 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 2
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9506726	A2	19950309	WO 1994-US9863	19940901
WO 9506726	A3	19950427		
W: BR, CA, CN, JP				
RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
EP 724635	A1	19960807	EP 1994-926657	19940901

R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LI, LU, MC, NL, PT, SE

CN 1133063	A	19961009	CN 1994-193763		19940901
JP 09502095	T2	19970304	JP 1994-508259		19940901
BR 9407527	A	19971111	BR 1994-7527		19940901
US 6008201	A	19991228	US 1995-464052		19950605
US 6214543	B1	20010410	US 1995-461002		19950605
PRIORITY APPLN. INFO.:			US 1993-118442	A	19930902
			WO 1994-US9863	W	19940901
			US 1995-392210	A3	19950222

AB A DNA mol. conferring on Mycobacterium tuberculosis an ability to enter mammalian cells and to survive within macrophages is isolated from M. tuberculosis. The protein encoded by this gene fragment is useful in vaccines to prevent infection by Mycobacterium tuberculosis, while the antibodies raised against this protein can be employed in passively immunizing those already infected by the organism. Both these proteins and antibodies may be utilized in diagnostic assays to detect Mycobacterium tuberculosis in tissue or bodily fluids. The protein of the present invention can be assocd. with various other therapeutic materials, for administration to mammals, particularly humans, to achieve uptake of those materials by such cells.

IT **148514-28-7**
RL: PRP (Properties)
(nucleotide sequence; cloning of gene for cellular uptake of Mycobacterium tuberculosis)

FILE 'HOME' ENTERED AT 10:46:31 ON 03 APR 2002

```
=> fil reg; d que l2
FILE 'REGISTRY' ENTERED AT 10:45:21 ON 03 APR 2002
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2002 American Chemical Society (ACS)

STRUCTURE FILE UPDATES:   31 MAR 2002  HIGHEST RN 403694-27-9
DICTIONARY FILE UPDATES:  31 MAR 2002  HIGHEST RN 403694-27-9

TSCA INFORMATION NOW CURRENT THROUGH July 7, 2001

  Please note that search-term pricing does apply when
  conducting SmartSELECT searches.

Crossover limits have been increased.  See HELP CROSSOVER for details.

Calculated physical property data is now available.  See HELP PROPERTIES
for more information.  See STNote 27, Searching Properties in the CAS
Registry File, for complete details:
http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf

The P indicator for Preparations was not generated for all of the
CAS Registry Numbers that were added to the H/Z/CA/CAplus files between
12/27/01 and 1/23/02.  Use of the P indicator in online and SDI searches
during this period, either directly appended to a CAS Registry Number
or by qualifying an L-number with /P, may have yielded incomplete results.
As of 1/23/02, the situation has been resolved.  Also, note that searches
conducted using the PREP role indicator were not affected.

Customers running searches and/or SDIs in the H/Z/CA/CAplus files
incorporating CAS Registry Numbers with the P indicator between 12/27/01
and 1/23/02, are encouraged to re-run these strategies.  Contact the
CAS Help Desk at 1-800-848-6533 in North America or 1-614-447-3698,
worldwide, or send an e-mail to help@cas.org for further assistance or to
receive a credit for any duplicate searches.

L2            23 SEA FILE=REGISTRY ABB=ON  ACATCAAAGTGATTCGCG|CGCGAATCACTTTGATGT
                 |CATGCCGTCGTATTGCTG|CAGCAATACGACGGCATG/SQSN
```

Seq 384 and their complements

```
=> d rn cn sql kwic nte l2 1-23; fil capl; s l2

L2   ANSWER 1 OF 23  REGISTRY  COPYRIGHT 2002 ACS
RN   391849-09-5  REGISTRY
CN   GenBank Z95390 (9CI)  (CA INDEX NAME)
SQL  43401

SEQ  15851 caaccagcaa tacgacggca tgtcacggct aagtggctac ctgacccccc
                ====== ========== ==
HITS AT:   15855-15872

L2   ANSWER 2 OF 23  REGISTRY  COPYRIGHT 2002 ACS
RN   385150-46-9  REGISTRY
CN   GenBank AR147696 (9CI)  (CA INDEX NAME)
SQL  650

SEQ    151 cgcggcatca ccctgagcaa ccagcaatac gacggcatgt cacggctaag
                               ========= =========
HITS AT:   172-189

L2   ANSWER 3 OF 23  REGISTRY  COPYRIGHT 2002 ACS
RN   385150-45-8  REGISTRY
```

sequence length

CN GenBank AR096715 (9CI) (CA INDEX NAME)
SQL 650

SEQ 151 cgcggcatca ccctgagcaa ccagcaatac gacggcatgt cacggctaag
========= =========
HITS AT: 172-189

L2 ANSWER 4 OF 23 REGISTRY COPYRIGHT 2002 ACS
RN 384520-79-0 REGISTRY
CN GenBank AR147694 (9CI) (CA INDEX NAME)
SQL 1535

SEQ 1051 agcaaccagc aatacgacgg catgtcacgg ctaagtggct acctgacccc
==== ========== ====
HITS AT: 1057-1074

L2 ANSWER 5 OF 23 REGISTRY COPYRIGHT 2002 ACS
RN 384520-78-9 REGISTRY
CN GenBank AR096713 (9CI) (CA INDEX NAME)
SQL 1535

SEQ 1051 agcaaccagc aatacgacgg catgtcacgg ctaagtggct acctgacccc
==== ========== ====
HITS AT: 1057-1074

L2 ANSWER 6 OF 23 REGISTRY COPYRIGHT 2002 ACS
RN 362724-95-6 REGISTRY
CN DNA (Mycobacterium tuberculosis strain CDC1551) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 1: PN: US6294328 SEQID: 2 claimed DNA 09/103840
SQL 4403765

Alignment ~~Results~~ appears in Compugen .rni results (highlighted in purple)

SEQ **DISPLAY LENGTH EXCEEDS SYSTEM LIMITS**

L2 ANSWER 7 OF 23 REGISTRY COPYRIGHT 2002 ACS
RN 362640-82-2 REGISTRY
CN DNA (Mycobacterium tuberculosis strain H37Rv) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 1: PN: US6294328 SEQID: 1 claimed DNA 09/103840
SQL 4411529

Alignment appears in Compugen .rni results (highlighted in purple)

SEQ
*** DISPLAY LENGTH EXCEEDS SYSTEM LIMITS ***

L2 ANSWER 8 OF 23 REGISTRY COPYRIGHT 2002 ACS
RN 362070-38-0 REGISTRY
CN DNA, d(C-A-T-G-C-C-G-T-C-G-T-A-T-T-G-C-T-G) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 4: PN: US20010023065 SEQID: 4 claimed DNA
SQL 18

SEQ 1 catgccgtcg tattgctg
========== ========
HITS AT: 1-18

L2 ANSWER 9 OF 23 REGISTRY COPYRIGHT 2002 ACS
RN 362070-37-9 REGISTRY
CN DNA, d(A-C-A-T-C-A-A-A-G-T-G-A-T-T-C-G-C-G) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 3: PN: US20010023065 SEQID: 3 claimed DNA
SQL 18

```
SEQ      1 acatcaaagt gattcgcg
           ========== ========
HITS AT:   1-18

L2   ANSWER 10 OF 23  REGISTRY  COPYRIGHT 2002 ACS
RN   362070-36-8  REGISTRY
CN   DNA (Mycobacterium tuberculosis strain H37Rv repetitive DNA REP13E12)
     (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   2: PN: US20010023065 SEQID: 2 claimed DNA
SQL  453

SEQ      1 gatcggcgag gcgcacatca aagtgattcg cgcccttttt cgcccatttg
                     ====== ========== ==
       201 ccgcaaacgc gcatcaccct gagcaaccag caatacgacg gcatgttcacg
                                    === ========== =====
HITS AT:   15-32, 228-245

L2   ANSWER 11 OF 23  REGISTRY  COPYRIGHT 2002 ACS
RN   362070-35-7  REGISTRY
CN   DNA (Mycobacterium tuberculosis strain H37Rv repetitive DNA REP13E12
     region-containing fragment) (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   1: PN: US20010023065 SEQID: 1 claimed DNA
SQL  1393

SEQ    601 aaccagcaat acgacggcat gtcacggcta agtggctacc tgacccccca
            ======= ========== =
HITS AT:   604-621

L2   ANSWER 12 OF 23  REGISTRY  COPYRIGHT 2002 ACS
RN   335512-15-7  REGISTRY
CN   GenBank AE007028 (9CI)  (CA INDEX NAME)
SQL  17783

SEQ    851 gggggtcagg tagccactta gccgtgacat gccgtcgtat tgctggttgc
                                   === ========== =====
      1051 gggtggacac atccaccgcg gcgggcaggt gggcgaaaaa gggcgcgaat
                                                     =======
      1101 cactttgatg tgcgcctcgc cgatcaggcc ctggcgttgg gcggtggcgg
           ========== =
HITS AT:   878-895, 1094-1111
NTE  doublestranded

L2   ANSWER 13 OF 23  REGISTRY  COPYRIGHT 2002 ACS
RN   335511-08-5  REGISTRY
CN   GenBank AE006921 (9CI)  (CA INDEX NAME)
SQL  9764

SEQ   9001 gacatgccgt cgtattgctg gttgctcagg gtgatgccgc gtttgcgggc
           ======== ==========
      9201 caggtgggcg aaaaagggcg cgaatcactt tgatgtgcgc ctcgccgatc
                             == ========== ======
HITS AT:   9003-9020, 9219-9236
NTE  doublestranded

L2   ANSWER 14 OF 23  REGISTRY  COPYRIGHT 2002 ACS
RN   209154-53-0  REGISTRY
CN   GenBank I86264 (9CI)  (CA INDEX NAME)
SQL  12412

SEQ    451 ctgatcggcg aaggcgcaca tcaaagtgat tcgcgccctt tttcggccca
```

```
                              === ========== =====
       651 cgaacgcgcc cgcaaacgcg gcatcaccct gagcaaccag caatacgacg
                                                  === ==========
       701 gcatgtcacg gctaagtggc tacctgaccc cccaagcgcg ggccaccttt
           =====
HITS AT:   468-485, 688-705

L2   ANSWER 15 OF 23  REGISTRY  COPYRIGHT 2002 ACS
RN   202636-85-9  REGISTRY
CN   GenBank AF041819 (9CI)  (CA INDEX NAME)
SQL  10019

SEQ   6351 cgcgcttggg gggtcaggta gccacttagc cgtgacatgc cgtcgtattg
                                                 ===== ==========
      6401 ctggttgctc agggtgatgc cgcgtttgcg ggcgcgttcg gtgtcggtga
           ===
      6601 gcgcgaatca ctttgatgtg cgcctcgccg atcaggccct ggcgttgggc
            ========= =========
HITS AT:   6386-6403, 6602-6619
NTE  doublestranded

L2   ANSWER 16 OF 23  REGISTRY  COPYRIGHT 2002 ACS
RN   190741-74-3  REGISTRY
CN   DNA (Mycobacterium tuberculosis strain H37Rv 32,437-nucleotide fragment)
     (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   GenBank AL123456
CN   GenBank Z95586
SQL  32437

SEQ  17601 cacttagccg tgacatgccg tcgtattgct ggttgctcag ggtgatgccg
                        ======= ========== =
HITS AT:   17614-17631
NTE  doublestranded

L2   ANSWER 17 OF 23  REGISTRY  COPYRIGHT 2002 ACS
RN   190550-78-8  REGISTRY
CN   DNA (Mycobacterium tuberculosis strain H37Rv 43,401-nucleotide fragment)
     (9CI)  (CA INDEX NAME)
OTHER NAMES:
CN   16: PN: WO0102555 FIGURE: 12A unclaimed DNA
SQL  42808

SEQ  15201 ccgacaccga acgcgcccgc aaacgcggca tcaccctgag caaccagcaa
                                                       ======
     15251 tacgacggca tgtcacggct aagtggctac ctgacccccc aagcgcgggc
           ========== ==
HITS AT:   15245-15262

L2   ANSWER 18 OF 23  REGISTRY  COPYRIGHT 2002 ACS
RN   182239-95-8  REGISTRY
CN   DNA (Mycobacterium tuberculosis 216-amino acid mammalian cell invasion
     protein gene) (9CI)  (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN   Deoxyribonucleic acid (Mycobacterium tuberculosis 216-amino acid mammalian
     cell invasion protein gene)
SQL  650

SEQ    151 cgcggcatca ccctgagcaa ccagcaatac gacggcatgt cacggctaag
                                 ========= =========
HITS AT:   172-189

L2   ANSWER 19 OF 23  REGISTRY  COPYRIGHT 2002 ACS
```

RN 178353-85-0 REGISTRY
CN DNA (Mycobacterium tuberculosis strain H37Rv 38,380-nucleotide fragment) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AL123456
CN GenBank Z74410
SQL 36198

SEQ 12451 tagccactta gccgtgacat gccgtcgtat tgctggttgc tcagggtgat
=== ========== =====
HITS AT: 12468-12485
NTE doublestranded

L2 ANSWER 20 OF 23 REGISTRY COPYRIGHT 2002 ACS
RN 176456-75-0 REGISTRY
CN GenBank U43540 (9CI) (CA INDEX NAME)
SQL 2334

SEQ 2001 gcaaacggca tcaccctgag caaccagcaa tacgacggca tgtcacggct
====== ========== ==
HITS AT: 2025-2042
NTE doublestranded

L2 ANSWER 21 OF 23 REGISTRY COPYRIGHT 2002 ACS
RN 170316-55-9 REGISTRY
CN DNA (Mycobacterium bovis strain ATCC 19210 deleted region-containing 1604-nucleotide fragment) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Deoxyribonucleic acid (Mycobacterium bovis strain ATCC 19210 deleted region-containing 1604-nucleotide fragment)
OTHER NAMES:
CN Deoxyribonucleic acid (Mycobacterium bovis virulent strain BCG 1604-nucleotide fragment)
CN GenBank U35021
SQL 1604

SEQ 451 gatcggcgag gcgcacatca aagtgattcg cgccctttt cgcccacctg
====== ========== ==
651 gcccgcaaac gcggcatcac cctgagcaac cagcaatacg acggcatgtc
========== ========
HITS AT: 465-482, 681-698
NTE doublestranded

L2 ANSWER 22 OF 23 REGISTRY COPYRIGHT 2002 ACS
RN 170316-51-5 REGISTRY
CN DNA (Mycobacterium bovis strain BCG clone pGM542 deleted region-containing 9281-nucleotide fragment) (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN Deoxyribonucleic acid (Mycobacterium bovis strain BCG clone pGM542 deleted region-containing 9281-nucleotide fragment)
OTHER NAMES:
CN Deoxyribonucleic acid (Mycobacterium bovis strain BCG clone pGM542 R3 9281-nucleotide fragment)
CN GenBank U35017
SQL 9281

SEQ 451 gatcggcgag gcgcacatca aagtgattcg cgccctttt cgcccacctg
====== ========== ==
651 gcccgcaaac gcggcatcac cctgagcaac cagcaatacg acggcatgtc
========== ========
HITS AT: 465-482, 681-698
NTE doublestranded